

# A Hierarchical Nonparametric Bayesian Model that Integrates Multiple Sources of Lifetime Information to Model Large-Scale System Reliability

Richard L. Warr \*

Department of Mathematics and Statistics  
Air Force Institute of Technology  
Wright-Patterson Air Force Base, OH 45433  
(richard.warr@afit.edu)

and

Brandon M. Greenwell \*

Department of Mathematics and Statistics  
Air Force Institute of Technology  
Wright-Patterson Air Force Base, OH 45433  
(brandon.greenwell@afit.edu)

December 16, 2014

## Abstract

The need for new large-scale reliability models is becoming apparent as the amount of available data is expanding at a dramatic rate. Often complex systems have thousands of components. Each of these components and their respective subsystems could have many, few, or no test data. The large number of components creates a massive estimation project that challenges the computational feasibility of traditional reliability models. The solution presented in this work suggests a hierarchical nonparametric Bayesian framework, using beta-Stacy processes. In this Bayesian framework, time-to-event distributions are estimated from sample data (which may be randomly right censored), and possible expert opinion. These estimates can be used to compute and predict system reliability.

---

\*The views expressed in this article are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

KEY WORDS: BETA-STACY PROCESS, BIG DATA, COMPLEX SYSTEMS, DIRICHLET PROCESS, RIGHT CENSORING

## 1 Introduction

As the complexity or size of systems continues to increase, implementation of traditional reliability models becomes impractical. The “big data” age is here. Modelers tend to transform these big data to conform with well-known techniques; this approach is not optimal. Instead, new models are needed that are built for big data problems. These new models should be scalable, flexible, and have reasonable computation time.

Reliability modeling of large systems has been problematic. In many instances the components or small subsystems are relatively easy to test, and a large amount of data are collected at these levels. The larger subsystems and system, in general, become harder and/or more expensive to test. An effective system reliability model should incorporate as much of the component, subsystem, and system data as possible.

Many modern systems have large numbers of components. The typically high cost of full system data drives the need for incorporating available component and subsystem data into the system reliability estimate. A strictly empirical result is not well suited for combining multi-level reliability data. Moreover, a parametric approach could include hundreds to thousands of parametric model assumptions. If incorrect, these assumptions may severely distort the reliability estimate.

The approach taken in this article is based on the concepts of Warr and Collins (2014). The model presented there cannot incorporate censored data which is unrealistic for reliability modeling. This article presents a model that can incorporate right-censored data using similar concepts. We avoid a parametric modeling approach for two primary reasons. First, parametric families must be assumed for lifetimes. If this is done for each component and subsystem, the model assumptions become too unwieldy to validate, when the number of components in the system is large. Another reason to avoid a parametric framework is to avoid inference in a large parameter space.

Bayesian nonparametric models have excellent qualities for this problem. Some Bayesian nonparametric models contain “conjugate” properties that make computation quite simple. Another quality they have is the flexibility to model any CDF shape. Finally, Bayesian nonparametric models allow for expert opinion to be included in the analysis.

The paper develops a system model that can incorporate expert opinion and component/subsystem data. This model is adaptable to large scale systems, has reasonable computation time, and allows for a full spectrum of expert opinion (from none to a significant amount).

## 1.1 Background

In reliability analysis, it is often the goal to estimate an unknown CDF. A key characteristic of reliability data, however, is the presence of censored observations. Quite often, we are interested in the reliability of a system that is composed of several smaller components, each with their own distribution function describing the probability of failure at any given point in time. Let  $F(t)$  and  $R(t) = 1 - F(t)$  denote the distribution function and reliability function of a random variable, respectively.

Much of the early work on system reliability involved binomial data (e.g., the components either passed or failed inspection), as in Cole (1975). Other important works include Mastran (1976) and Mastran and Singpurwalla (1978). Reese et al. (2011) propose a method for assessing the reliability of complex multicomponent systems using Markov Chain Monte Carlo (MCMC) methods. The Bayesian nature of their method easily allows for the pooling of information across similar components, and the incorporation of expert opinion. Their work is further extended in Guo and Wilson (2013) to binary, lifetime, degradation, and expert opinion data at any level of the system. Warr and Collins (2014) use the *Dirichlet process* (DP) to make a nonparametric assessment of the reliability of multicomponent systems. A drawback of this approach, however, is that it assumes the data are not censored. In this paper we extend the methodology to allow for right-censored data by replacing the Dirichlet process with the more general *Beta-Stacy process* (Walker and Muliere; 1997). Extensive literature exists on system reliability assessment from both frequentist and Bayesian perspectives. For a thorough overview on this literature, see Reese et al. (2011) and Guo and Wilson (2013).

If we know little about  $F(t)$  (or equivalently  $R(t)$ ), then it makes sense to adopt a non-parametric approach for making inference and predictions. Within the frequentist framework, this is typically done by estimating  $F(t)$  using the empirical distribution function, or in the case of censored data, the Kaplan-Meier estimator (Kaplan and Meier; 1958). However, in the Bayesian framework, we specify a prior process (e.g., a DP) on the space of all distribution functions  $\mathcal{F}$ . Doksum (1974) introduced the *neutral to the right process* (NRP), a special class of random distribution functions. An important feature of NRPs is: given an NRP prior and data (which may be randomly right censored) the posterior is also a NRP. In this sense, a NRP prior is *structurally conjugate*. Susarla and Van Ryzin (1976) use a Dirichlet process (DP) prior for the unknown distribution function  $F(t)$ . However, the DP may not provide an adequate prior for  $F(t)$  in the case of lifetime data. Ferguson and Phadia (1979) extend their results to a more general class of prior distributions for  $F(t)$  which can incorporate right-censored observations, namely, the NRP prior. Walker and Muliere (1997) introduce the Beta-Stacy process (BSP) prior, a neutral to the right process prior that is structurally and parametrically conjugate. The authors also show the BSP as a generalization of the Dirichlet process. In fact, given right-censored observations and a DP prior for  $F(t)$ , the posterior of  $F(t)$  is a Beta-Stacy process. Thus, the BSP provides a convenient prior for the space of distribution functions.

## 1.2 The Beta-Stacy Process

The Beta-Stacy process (BSP) is a generalization of the Dirichlet Process. It is one type of NRP. BSPs can be used to model an unknown CDF  $F(t)$ . From the Bayesian perspective our belief/knowledge of  $F(t)$  is considered to be random. As we collect additional data about  $F(t)$  theoretically our belief about  $F(t)$  should approach the actual function.

One distinguishing difference between the BSP and the DP is that the precision parameter in the BSP is a function which can vary over the support, whereas the precision parameter of the DP is a constant. This variable precision parameter in the BSP provides enough flexibility to be “conjugate” with randomly right censored data. The DP is only considered to be conjugate with complete data (i.e., for data where the exact failure time is known). A BSP is defined by two functions. The first is denoted as  $\alpha(t)$ ; we call this the precision parameter. The precision parameter must be positive; the larger the precision the more certain the belief of  $F(t)$ . The second parameter is denoted by  $G(t)$  and called the base measure. The base measure is a proper CDF and is also the expected value of the BSP. To illustrate, for a given  $t$ , if

$$F(t) \sim BSP(\alpha(t), G(t)), \text{ then } E[F(t)] = G(t). \quad (1)$$

In this paper  $G(t)$  will always be the CDF of a discrete random variable. This does not limit or preclude  $F$  from being a continuous CDF, it merely limits where an estimate for  $F$  is obtained.

For the sake of simplicity we avoid discussing how the BSP is formally defined by a Lévy process. However, for this model it is necessary to know the first two moments of  $F(t) \sim BSP(\alpha(t), G(t))$ . The calculations are greatly simplified with the current restriction that  $G(t)$  is the CDF of a positive discrete random variable. The first moment is defined in Equation 1 and is derived in Walker and Muliere (1997). The second moment is derived from a property of the jumps in the Lévy process. The distribution of the jumps can be transformed to a beta distribution and the jumps of the Lévy process can be related back to  $F(t)$  (which is a BSP). Let  $t_1 < t_2 < \dots < t_n$  be all the time increments such that  $G$  is discontinuous at each (i.e.,

$$\lim_{t \rightarrow t_k^-} G(t) \neq \lim_{t \rightarrow t_k^+} G(t) \quad (2)$$

where  $k \in \{1, 2, \dots, n\}$ ). Additionally, define  $t_0 \equiv 0$  and  $t_{n+1} \equiv \infty$ . For an arbitrary  $t > 0$  let  $m \in \{0, 1, 2, \dots, n\}$  be the index such that  $t_m \leq t < t_{m+1}$  then

$$E[(F(t))^2] = \left( \prod_{i=0}^m \frac{(1 - G(t_i))[\alpha(t_i)(1 - G(t_i)) + 1]}{(1 - G(t_i-))[\alpha(t_i)(1 - G(t_i-)) + 1]} \right) - 1 + 2G(t). \quad (3)$$

Where

$$G(t_i-) = \lim_{t \rightarrow t_i^-} G(t). \quad (4)$$

At this point we should mention that the theory for BSPs with discrete base measures only applies at the jumps  $t_k$ . So the statements in Equations 1 and 3 are only formally defined for  $t \in \{t_1, t_2, \dots, t_{n-1}\}$ . From an intuitive perspective, if  $t_1 \leq t < t_n$  then treating  $t$  as a continuous variable poses no problems and provides a consistent estimator, much the same as the ECDF. Although the estimates are at discrete points, assuming  $F(t)$  is continuous makes sense, because one can only guess at the value of  $F(t)$  where data are observed (nonparametric estimation). For example, if the posterior distribution of  $F(t_1)$  (for fixed  $t_1$ ) is centered at 0.2 with small uncertainty, it follows that  $P(T \leq t_1)$  is near 0.2. However, for  $0 < t^* < t_1$ , no information exists about the value of  $F(t^*)$ , other than  $F(t^*) \leq F(t_1)$ . This approach does not force  $F(t)$  to be discrete, but the theoretical estimates are defined at discrete points. Also, for estimates of  $F(t)$  where the base measure is 0 or 1, 0 and 1 will be the respective estimates, with a variance which does not exist. This reiterates that where there is no data (or information), nonparametric estimation does not inform. With these concepts in mind the posteriors for BSP priors are introduced.

Now assume prior information about a CDF is contained in the prior  $F(t) \sim \text{BSP}(\alpha(t), G(t))$  and data from the random variable  $T_i \sim F$  are collected. The purpose for using the BSP in this setting is to allow for right censoring, therefore  $T$  (the data vector) may include right censored data which is denoted in the corresponding vector  $C$ . If  $C_1 = 1$  then  $T_1$  is not censored and if  $C_1 = 0$  then  $T_1$  is censored. Define

$$M(t) = \sum_{i=1}^m I(T_i \geq t), \text{ (number of units not failed just before time } t) \text{ and}$$

$$J(t) = \sum_{i=1}^m C_i I(T_i = t), \text{ (number of failures that occurred at time } t).$$

The posterior  $F(t)|T$ , as defined in Walker and Muliere (1997), is another BSP with new base measure

$$G^*(t) = 1 - \prod_{i=1}^m 1 - \frac{\alpha(t_i)(G(t_i) - G(t_i-)) + J(t_i)}{\alpha(t_i)(1 - G(t_i-)) + M(t_i)} \quad (5)$$

and new precision parameter

$$\alpha^*(t) = \frac{\alpha(t)(1 - G(t)) + M(t) - J(t)}{1 - G^*(t)}. \quad (6)$$

To demonstrate Equations 5 and 6 consider the CDF  $H(t)$  which is defined as:

$$H(t) = \begin{cases} 0 & \text{for } t < 1 \\ 1/3 & \text{for } 1 \leq t < 2 \\ 2/3 & \text{for } 2 \leq t < 3 \\ 1 & \text{for } t \geq 3. \end{cases} \quad (7)$$

Let  $H(t)$  be the prior and suppose no other data is collected. Then the posterior  $G^*(t)$  is equal to the prior  $H(t)$ .

$$G^*(1) = 1 - \prod_{i=1}^1 \left\{ 1 - \frac{\alpha(t_i)(H(t_i) - H(t_i-)) + J(t_i)}{\alpha(t_i)(H(t_i) - H(t_i-)) + J(t_i)} \right\} = 1 - \left( 1 - \frac{\alpha(1)(1/3 - 0) + 0}{\alpha(1)(1 - 0) + 0} \right) = 1/3 = H(1),$$

$$G^*(2) = 1 - \left( \frac{2}{3} \right) \left( 1 - \frac{\alpha(2)(2/3 - 1/3) + 0}{\alpha(2)(1 - 1/3) + 0} \right) = 2/3 = H(2),$$

and

$$G^*(3) = 1 - \left( \frac{2}{3} \right) \left( \frac{1}{3} \right) \left( 1 - \frac{\alpha(3)(1 - 2/3) + 0}{\alpha(3)(1 - 2/3) + 0} \right) = 1 = H(3).$$

The posterior precision should also be equal to the prior precision and we obtain

$$\alpha^*(1) = \frac{\alpha(1)1 + 0 - 0}{1} = \alpha(1),$$

$$\alpha^*(2) = \frac{\alpha(2)2/3 + 0 - 0}{2/3} = \alpha(2),$$

and

$$\alpha^*(3) = \frac{\alpha(3)1/3 + 0 - 0}{1/3} = \alpha(3).$$

The converse scenario also produces reasonable results. Assume no prior information is available (i.e.,  $\alpha(t) \equiv 0$ ) and there are three failures which occur at time points 1, 2, and 3, then

$$G^*(1) = 1 - \prod_{i=1}^1 \left\{ 1 - \frac{0(H(t_i) - H(t_i-)) + J(t_i)}{0(H(t_i) - H(t_i-)) + J(t_i)} \right\} = 1 - \left( 1 - \frac{1}{3} \right) = 1/3,$$

$$G^*(2) = 1 - \left( \frac{2}{3} \right) \left( 1 - \frac{1}{2} \right) = 2/3,$$

and

$$G^*(3) = 1 - \left( \frac{2}{3} \right) \left( \frac{1}{3} \right) \left( 1 - \frac{1}{1} \right) = 1.$$

Which is just the empirical CDF. For the precision parameter we have

$$\alpha^*(1) = \frac{0(1 - G(1)) + 3 - 1}{1 - 1/3} = 3,$$

$$\alpha^*(2) = \frac{0(1 - G(2)) + 2 - 1}{1 - 2/3} = 3,$$

and

$$\alpha^*(3-) = \frac{0(1 - G(3-)) + 1 - 0}{1 - 2/3} = 3.$$

Which is a constant of 3 so the posterior is a DP as one would expect.

These basic formulas and properties of BSPs are necessary to build the proposed model.

## 2 Methodology

This section describes how the proposed system reliability model is assembled. After the model is presented a few other topics are addressed to aid in using the model.

### 2.1 Hierarchical Model

The goal of the model is to use the information of the components as a prior for the subsystem. The basic concept is quite simple and can be recursively applied to a system with many components. The process begins first at the lowest level; the components are modeled using a BSP (they may or may not have a BSP prior). Next the components for a particular subsystem are “fused” (or aggregated) together to make a BSP prior for the subsystem. Additional prior information about a subsystem can be used to update this prior. Then this BSP prior is updated with subsystem data to obtain a BSP posterior for the subsystem. This process is iterated, treating the subsystems as components at the next level in the system hierarchy.

To model this we propose the following steps:

- For each component choose a discrete BSP prior, or if no prior information, set its precision parameter to zero.
- For each component, collect data and update its prior to obtain a posterior BSP.
- Compute the first and second moments from the BSP posterior for each component. These moments are functions, so for a particular  $t$  we have a first and second moment for each component.
- Calculate the first and second moments of the merged components

Once the following steps have been taken, we use the merged component’s information (first and second moment functions) and fit it to a BSP prior, which is a prior for the subsystem. This follows a similar method as Warr and Collins (2014), but now the precision is also preserved in addition to the mean.

Because the BSP has only two “parameters,” having the first and second moments at each time  $t$  is enough information to find a BSP with those same moments at time  $t$ .

The remainder of this section discusses how to find the first and second moments for the combined components after which the process for finding the base measure and precision is shown. A fundamental model assumption is that each component is independent of one another. This allows straight forward calculations of combinations of components. To complete this task, one must have a reliability block diagram of the subsystem which shows how the components are interrelated. The two basic relations of *in series* or *in parallel* determine how the information is combined.

Finding the prior mean of the subsystem using component data is demonstrated in Warr and Collins (2014). Essentially, for two components in parallel their combined information

is contained in the maximum of the two random variables. In Figure 1a the subsystem CDF is:

$$F_S(t) = P(X_1 \leq t \text{ and } X_2 \leq t) = P(X_1 \leq t)P(X_2 \leq t) = F_{X_1}(t)F_{X_2}(t).$$

Basically the two base measures of the components are multiplied together to obtain the base measure for the subsystem. Therefore

$$G_S(t) \equiv E[F_S(t)] = E[F_{X_1}(t)F_{X_2}(t)] = E[F_{X_1}(t)]E[F_{X_2}(t)] = G_1(t)G_2(t).$$

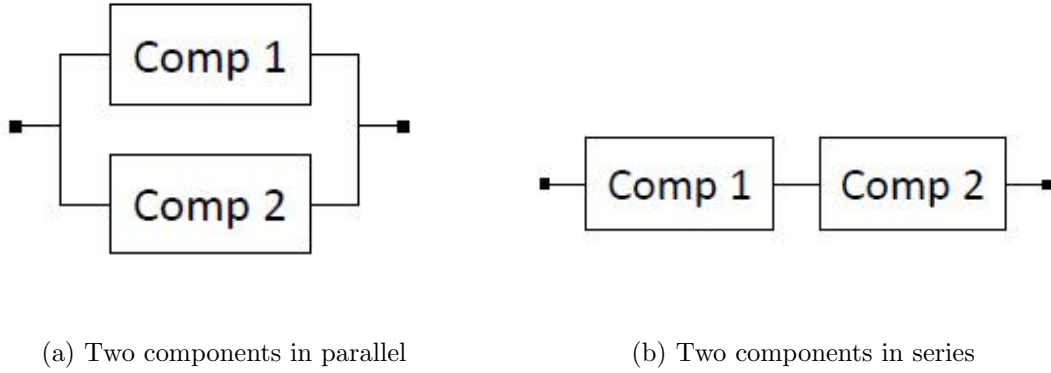


Figure 1: Two component systems

If two components are in series (as in Figure 1b) the subsystem CDF is:

$$F_S(t) = P(X_1 \leq t \text{ or } X_2 \leq t) = 1 - P(X_1 > t \text{ and } X_2 > t) = 1 - P(X_1 > t)P(X_2 > t) = 1 - R_{X_1}(t)R_{X_2}(t).$$

Again  $R(t) \equiv 1 - F(t)$ , therefore

$$G_S(t) \equiv E[F_S(t)] = E[1 - R_{X_1}(t)R_{X_2}(t)] = 1 - E[1 - F_{X_1}(t)]E[1 - F_{X_2}(t)] = G_1(t) + G_2(t) - G_1(t)G_2(t).$$

Finding the second moment for the subsystem prior is also fairly straight forward. For the components in parallel the second moment of the subsystem is:

$$E[(F_S(t))^2] = E[(F_{X_1}(t)F_{X_2}(t))^2] = E[(F_{X_1}(t))^2]E[(F_{X_2}(t))^2].$$

See Equation 3 to calculate the last two quantities of the previous equation. For the components in series, the second moment of the subsystem is:

$$\begin{aligned} E[(F_S(t))^2] &= E[(1 - R_{X_1}(t)R_{X_2}(t))^2] = 1 - 2E[R_{X_1}(t)R_{X_2}(t)] + E[(R_{X_1}(t)R_{X_2}(t))^2] \\ &= 1 - 2E[R_{X_1}(t)]E[R_{X_2}(t)] + E[(R_{X_1}(t))^2]E[(R_{X_2}(t))^2] \end{aligned}$$



$$= 1 - 2G_1(t)G_2(t) + (1 - 2G_1(t) + E[(F_{X_1}(t))^2])(1 - 2G_2(t) + E[(F_{X_2}(t))^2]).$$

Although not elegant, the quantities from this last equation are known from the components.

The previous steps provide a first and second moment for the prior. The resulting process of this combined component information is not a BSP, but it can be well approximated by one. Therefore we find the BSP that has the same first two moments. Although this procedure seems mathematically tedious it is numerically fast. The first moment of the prior information is just the base measure for the prior BSP. The precision for the prior BSP is not as simple. Recall the form of Equation 3 but replace  $E[(F(t))^2]$  with the second moment of the combined prior information and  $G(t)$  with the first moment. For any  $t$  the only unknowns in the equation are the  $\alpha(t_i)$ . Because the same answer should be obtained regardless of the value of  $t$ , the convenient value of  $t_i$  is chosen. With simple algebra, a formula for the precision parameter is as follows:

$$\alpha(t_i) = \frac{(E[(F_S(t_{i-1}))^2] + 1 - 2G(t_{i-1}))(1 - G(t_i)) - (E[(F_S(t_i))^2] + 1 - 2G(t_i))(1 - G(t_{i-1}))}{(E[(F_S(t_i))^2] + 1 - 2G(t_i))(1 - G(t_{i-1}))^2 - (E[(F_S(t_{i-1}))^2] + 1 - 2G(t_{i-1}))(1 - G(t_i))^2}. \quad (8)$$

Although formally it does not exist, for Equation 8 define  $E[(F_S(t_0))^2] \equiv 0$ .

Equation 8 provides the computation needed to define the BSP prior for a simple two component subsystem. Most subsystems are not composed of just two components. Regardless of the subsystem complexity, pairs of components can be combined using the equations above to conceptually model a new component. This procedure is repeated until the first two moments of the subsystem are known. For an example see Figure 2. At each step in the figure a new component is defined and the subsystem is simplified.

Given the base measure and precision of the subsystem, comprised of the components, a BSP prior for the subsystem is defined. Once the prior is defined and subsystem data are available, Equations 5 and 6 are applied to find the posterior for the subsystem.

The process can be iterated as necessary by treating the subsystem as a component in a larger subsystem. The importance of computational efficiency is vital as the complexity of the system is increased. One important consideration of this model is as the complexity of a system grows, the computation time does not increase at an exponential rate. Thus the model has promise for big data applications, or in other words very large systems.

## 2.2 Priors for Components

This section provides a brief discussion on how to quantify prior information in this model. Although the model does not require prior information for components, if prior information is available it can be incorporated into the analysis.

A prior is given for a component's unknown CDF, which we denote  $F(t)$ . Again,  $F(t)$  is the true time-to-failure CDF for that component. If  $F(t)$  is continuous, an ideal prior

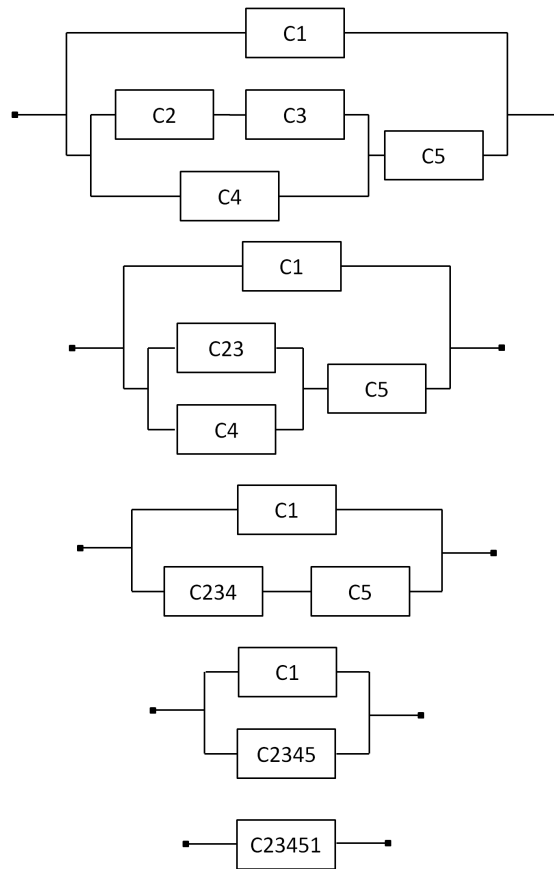


Figure 2: A step by step example of how to reduce a subsystem to a conceptually new component.

would also be continuous. However, prior information is often not ideal. Because the implementation of this model deals only with discrete CDFs, priors for  $F(t)$  should also be discrete.

One method to obtain a prior for  $F$  is to fit a DP to the modeler's prior information. To do this a modeler chooses  $l$  time points that can be quantified in a manner described below. Let  $t_0 \equiv 0$  and  $t_0 < t_1 < \dots < t_l < \infty$ . These time points can be arbitrary, but they should be chosen such that the modeler has some belief of the value of  $F(t_i)$ , for  $i \in \{1, 2, \dots, l\}$ . Once the belief of  $F$  is quantified at those time points a precision must be selected. The precision determines the certainty (or lack thereof) for the belief of  $F(t_i)$ . The DP's precision parameter can be interpreted as the amount of information contained in the prior. Therefore, if the precision is equal to 10, the prior DP has the weight of 10 posterior observations. Thus, if a modeler is quite certain (*a priori*) of the value of  $F(t)$ , the prior precision parameter should be large.

Although the DP prior is not the most flexible, it is fairly straight forward to implement. Clearly it would be possible to use a non constant precision parameter and have a BSP prior. This prior is more complex to implement, but can more accurately model prior information.

### 2.3 The Approximation of the Product of Betas with a Beta

A major provision of the proposed model is that a BSP can well approximate a process defined by two BSPs in series or parallel. This leads to the question, can the product of two beta random variables be well approximated by a different beta random variable? The answer to this question is yes.

For example, suppose we have four independent components in series with respective CDFs  $F_i(t)$ ,  $i = 1, 2, 3$ . The CDF for the system is then:

$$F(t) = 1 - \prod_{i=1}^3 [1 - F_i(t)].$$

If each  $F_i(t)$ , for a fixed  $t \geq 0$ , is distributed as  $Beta(a_i, b_i)$ , then what is the distribution for the system CDF  $F(t)$ ? Or in other words, what is the distribution of the product of independent beta random variables?

The outcome has a complicated distribution related to the Meijer G-function, see Springer and Thompson (1970) for additional details in the context of the product of betas random variables. However, this product can also be well approximated by a suitable two-parameter beta family:

$$f_X(x) = \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1} \cdot I(0 \leq x \leq 1), \quad a, b > 0,$$

where the shape parameters  $a$  and  $b$  are chosen by matching the first two moments or via maximum likelihood estimation.

For example, suppose that three beta random variables  $X_1$ ,  $X_2$ , and  $X_3$  have respective shape parameters  $(9, 3)$ ,  $(8, 3)$ , and  $(4, 2)$ . Springer and Thompson (1970) showed that the exact PDF for the product  $Y = X_1X_2X_3$  is

$$g_Y(y) = \frac{3,960}{7}y^3 - 1,980y^4 + 99,000y^7 + (374,220 + 356,400 \log y)y^8 - (443,520 - 237,600 \log y)y^9 - \frac{198,000}{7}y^{10}.$$

A histogram for a random sample of size 10,000 is given in Figure 3.

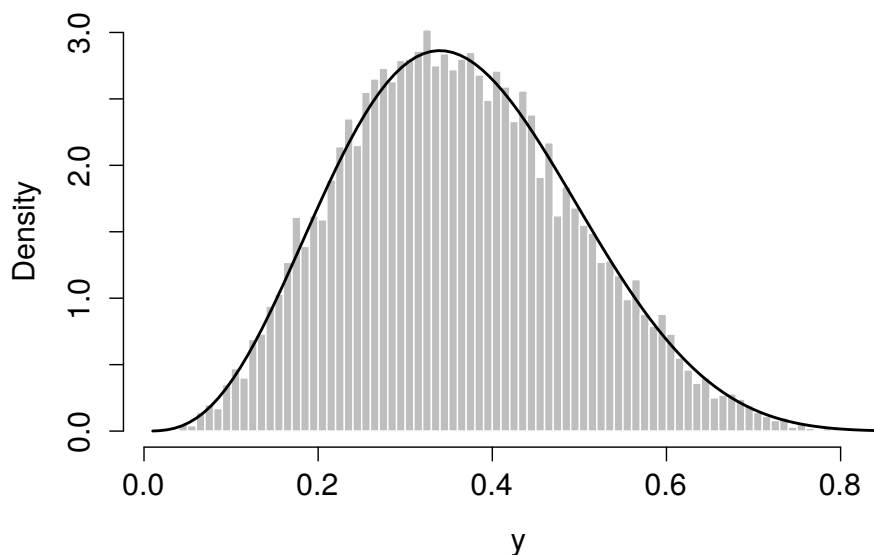


Figure 3: Data generated from the product of three independent beta random variables. The true density is shown with a histogram of 10,000 samples and the beta approximation is shown by a solid line.

This is only one example, but in principle it works well because the product of two random variables is a convolution in the log-transformed space. Convolution is a smoothing operation, so the resulting distribution of the product of two random variables is “better behaved” than the two original distributions. Thus the product of betas is well approximated by another beta, and this model is effective in capturing the behavior of a system’s time-to-event CDF. Although the product of two betas is not exactly another beta, this point-wise approximate distribution contains the necessary information/uncertainty of

$F(t)$ . For additional details on the approximation of the product of *iid* betas see Tang and Gupta (1984).

### 3 Application

To illustrate the application of this methodology, the CDF of the propulsion system in a small hybrid-electric remotely piloted aircraft (SHERPA) is modeled. SHERPA is a prototype remotely piloted aircraft (RPA) under development at the Air Force Institute of Technology (AFIT). SHERPA's desired mission profile involves takeoff, ingress to a target area, loitering over the target for surveillance, egress, and landing (Ausserer (2012)). RPAs powered by gasoline engines have a high power to weight ratio and are capable of long missions, but are noisy and thus less stealthy than desired. Electrically powered RPAs have a smaller acoustic signature, but battery weight is a challenge if electrical propulsion is to be used over an entire mission. A hybrid RPA uses both electrical and gasoline systems and possesses the ability to perform the primary purpose of the mission (loitering) solely on electric power.

Figure 4 shows a generic mission profile for SHERPA, along with the propulsion used. The gasoline engine (G) alone is used for ingress, egress, and landing; the electric motor (E) is used while loitering over the target; and both are employed (for maximum power) during takeoff and climb to cruising altitude. While cruising, the gas engine recharges the battery. For additional details refer to Warr and Collins (2014) for an analysis of SHERPA.

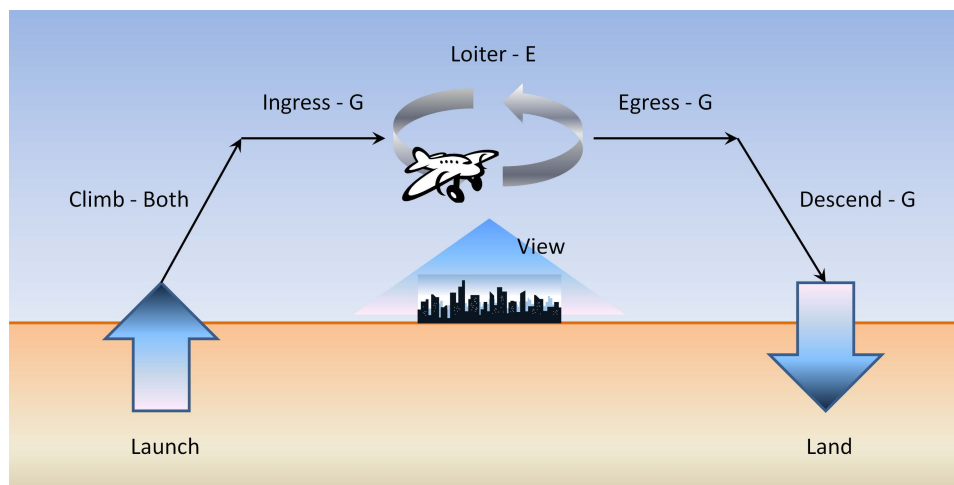


Figure 4: Mission profile for a small hybrid-electric remotely piloted aircraft

Table 1 lists the components of SHERPA's propulsion system. The interrelationship of components is shown by the reliability block diagram in Figure 5. The three common

Table 1: Component List for SHERPA Propulsion System

Electric Propulsion	Gasoline Propulsion	Common Parts
Motor	Engine	Propeller
Batteries	Gas Delivery	Drive shaft
Motor controller		Gearing
Serpentine belt		

components on the left are in series, since a failure of any one results in loss of propulsion. The two parallel branches to the right indicate that there is propulsion if either the gas engine or electric motor is functional. The reliability goal we consider is simply whether the aircraft can fly, though clearly its mission readiness would be compromised if only one propulsion mode were functional.

The data for this example are simulated to mimic the reliability of the components, subsystems, and system. Due to operational concerns actual test data are not available. No prior information is included for the components. Roughly 30 observations are simulated from each of the components, subsystems, and system. Random right censoring is introduced, where approximately 15% of the data are censored.

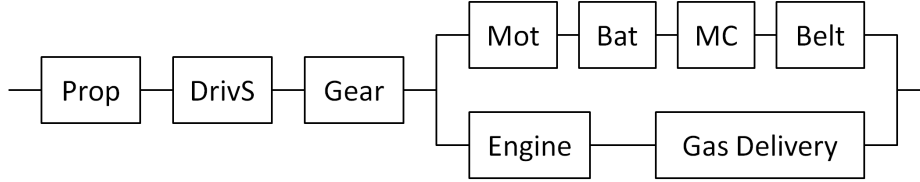


Figure 5: A reliability block diagram for the SHERPA propulsion system.

Although this model is simple in comparison to a large-scale system with hundreds or thousands of components, it demonstrates the capabilities of this methodology. The run-time for this example is approximate 2 seconds on a personal computer.

Figure 6 shows the resulting estimate of the system CDF. This posterior BSP provides the information to estimate or predict the reliability of this system. Contrast Figure 6 with Figure 7 which shows the estimate using only the full system test. This contrast shows the reduction of point-wise uncertainty for  $F(t)$  by incorporating component and subsystem data.

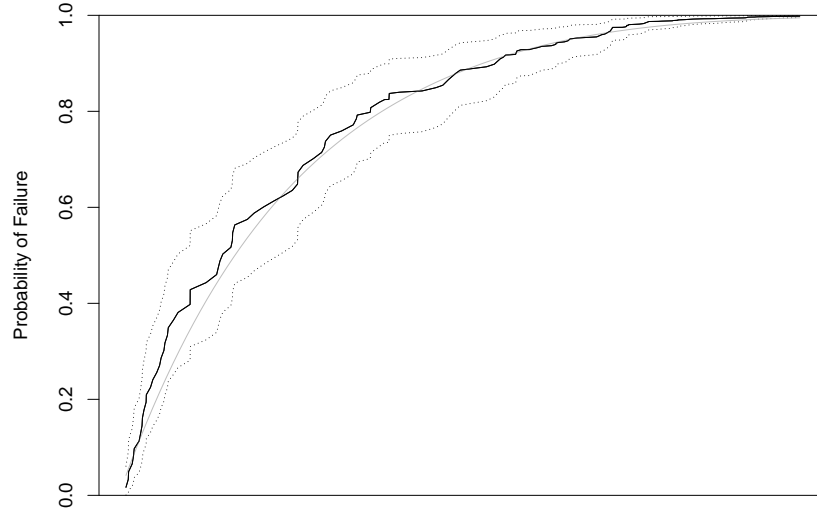


Figure 6: The estimated CDF,  $F(x)$ , in black. The dotted lines show the 95% point-wise probability intervals. The gray line is the true system CDF.

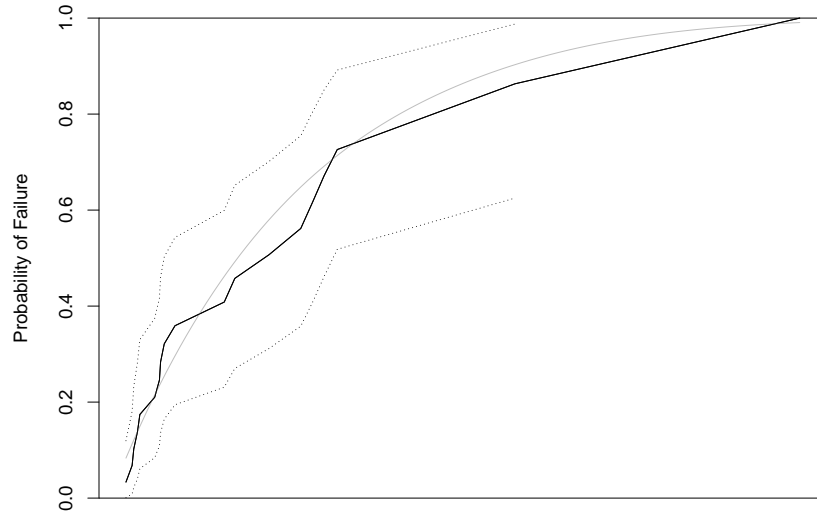


Figure 7: The estimated CDF,  $F(x)$ , in black using only the full system test data. The dotted lines show the 95% point-wise probability intervals. The gray line is the true system CDF.

## 4 Discussion

The method proposed in this paper is developed to model system reliability using system reliability data augmented with a prior BSP composed of component and subsystem data. Computation is very fast in relation to other methods, particularly Bayesian parametric methods that rely on MCMC computations.

One obvious model extension that might be explored is incorporating other types of reliability data such as interval censored data or pass/fail data. The proposed model is well suited for large-scale reliability applications.

## 5 Acknowledgments

The authors thank the B61 Life Extension Program at Los Alamos National Laboratory for providing funding which made this research possible.

## References

- Ausserer, J. K. (2012). *Integration, testing, and validation of a small hybrid-electric remotely-piloted aircraft*, Master's thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH.
- Cole, P. V. (1975). A Bayesian reliability assessment of complex systems for binomial sampling, *Reliability, IEEE Transactions on* **24**(2): 114–117.
- Doksum, K. A. (1974). Tailfree and neutral random probabilities and their posterior distributions, *The Annals of Probability* pp. 183–201.
- Ferguson, T. S. and Phadia, E. G. (1979). Bayesian nonparametric estimation based on censored data, *The Annals of Statistics* pp. 163–186.
- Guo, J. and Wilson, A. G. (2013). Bayesian methods for estimating system reliability using heterogeneous multilevel information, *Technometrics* **55**(4): 461–472.
- Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations, *Journal of the American statistical association* **53**(282): 457–481.
- Mastran, D. V. (1976). Incorporating component and system test data into the same assessment: a Bayesian approach, *Operations Research* **24**(3): 491–499.
- Mastran, D. V. and Singpurwalla, N. D. (1978). A Bayesian estimation of the reliability of coherent structures, *Operations Research* **26**(4): 663–672.



- Reese, S. C., Wilson, A. G., Guo, J., Hamada, M. S. and Johnson, V. E. (2011). A Bayesian model for integrating multiple sources of lifetime information in system-reliability assessments, *Journal of Quality Technology* **43**(2): 127–141.
- Springer, M. D. and Thompson, W. E. (1970). The distribution of products of beta, gamma and gaussian random variables, *SIAM Journal of Applied Mathematics* **18**(4): 721–737.
- Susarla, V. and Van Ryzin, J. (1976). Nonparametric Bayesian estimation of survival curves from incomplete observations, *Journal of the American Statistical Association* **71**(356): 897–902.
- Tang, J. and Gupta, A. (1984). On the distribution of the product of independent beta random variables, *Statistics & Probability Letters* **2**(3): 165–168.
- Walker, S. and Muliere, P. (1997). Beta-stacy processes and a generalization of the pólya-urn scheme, *The Annals of Statistics* pp. 1762–1780.
- Warr, R. L. and Collins, D. H. (2014). Bayesian nonparametric models for combining heterogeneous reliability data, *Journal of Risk and Reliability* **228**(2): 166–175.